DOCUMENT RESUME

ED 122 846 IR 003 471

AUTHOR Zipperer, W. C.
TITLE User Interface Models for Multidisciplinary Bibliographic Information Dissemination Centers.
PUB DATE Aug 75
NOTE 20p.; Paper presented at the American Chemical Society National Conference (Chicago, Illinois, August 24-29, 1975)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Data Bases; *Information Centers; Information Dissemination; *Information Retrieval; *Interaction Process Analysis; Interviews; Man Machine Systems; Models; Search Strategies; *Use Studies
IDENTIFIERS *Question Negotiation; University of California Los Angeles; University of Georgia

ABSTRACT

Two information dissemination centers at University of California at Los Angeles and University of Georgia studied the interactions between computer based search facilities and their users. The study, largely descriptive in nature, investigated the interaction processes between data base users and profile analysis or information specialists in constructing and revising search profiles. Data were gathered from users, intermediaries, and tape-recording search request interview sessions. Interview events were identified, labeled, and analyzed, and transition matrices were constructed for major groupings of events. On the basis of the data presearch and postsearch models of existing interfaces were constructed. The study indicated a need for data base information support, accommodation of a wide range of users, human intermediaries, built-in diagnostic capabilities, and alternative interaction mechanisms for the user. Results of the study showed that interface for the two centers was not linear, was highly adaptive, and had an important instructional component. Particularly significant was the finding that the human intermediary is an integral part of the interface essential to its adaptive capability. (Author/LS)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED122846

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

USER INTERFACE MODELS FOR MULTIDISCIPLINARY BIBLIOGRAPHIC INFORMATION DISSEMINATION CENTERS

In December of 1973, the two information dissemination centers at UCLA and the University of Georgia began a joint research project, under funding from the National Science Foundation, to study the interactions between computer based search facilities and users of these facilities. This study, which was completed early this year (1975), was largely descriptive in nature and was designed to investigate in some detail the reference process, particularly the interaction processes which occur between the users and the profile analysts or information specialists in constructing and revising search profiles to obtain satisfactory retrieval performance. Although the profile development process is probably the most critical aspect of the retrieval mechanism, very little information was available as to what actually takes place in this human information processing activity. While the data obtained in this study does, to a large part confirm intuition and common sense, it also goes further to provide quantitative data and, to some extent, insight into causal relationships. As a result of this study, a general model has been proposed of the retrieval process as it occurs in two centers using human intermediaries (that is, profile analysts, information specialists, and reference librarians) and this model differs considerably from the linear models found in most profile coding manuals. The implications of this model of the human profile coding process have also been projected to the networking environment.

Although the UCLA and Georgia research teams proceeded independently throughout the study, the basic data collection instruments were jointly defined, standardized, and tested prior to implementation in each of the centers. The same data collection forms were used at both sites wherever possible in order to identify center dependent characteristics which would

IR003471

have implications for a networking environment.

The approach taken was to collect data from the user and from the intermediary, and also to tape record the interview session during which the search request was negotiated into a search profile. These data along with the original and revised search profiles comprised the input data to the various subtasks designed to study the interface. At this point, I would like to refer you to the written report for discussions of methods and techniques so that this time can be spent on the results of the study. The final report which is about 300 pages long is available from the National Technical Information Service in Springfield, Virginia. I would like to use this time to concentrate on the actual models, their development, and their implications.

In developing the models, the tape recordings of the user-intermediary interviews were used by both research teams to conduct an event analysis of the interviews. The events that took place and the order in which they took place, were analyzed. At Georgia, investigation of the events taking place during the interview was done by doing a time-and-sequence analysis. The sequence of events were identified along with the time segment in which they occurred. Fifty-one unique events were identified from the interviews and these were grouped into nine categories. The first slide shows seven of the principal major categories, the other two being Administrative Procedures which generally included events such as getting the user's name and address; and

QUESTION NEGOTIATION

PROFILE DEVELOPMENT

TUTORIAL ACTIVITIES

SEARCH TYPE SELECTION

STRATEGY FORMULATION

SYSTEM DESCRIPTION

DATA BASE SELECTION

Slide 1

Diversionary Activities which includes external interruptions of the interview, such as phone calls. Analysis of the mean times spent in each of these seven categories shows a decrease in the order presented on the slide. The first category, Question Negotiation, includes events related to presentation and clarification of the user's question. Profile Development includes a wide range of vocabulary related events, such as selection, deletion, and truncation of search terms. Tutorial activities refer to detailed explanations of such things as data base indexing policies, retrospective holdings, and data base coverage. Search Type Selection refers to the choice of current awareness of retrospective searches or both. Strategy Formulation refers to events concerned with analysis of the questions, identification of the concepts, and specification of the relationships between concepts. System Description corresponds to descriptions of general system characteristics, such as search schedules and procedures; and Data Base Selection is self explanatory. Question Negotiation, Profile Development, and the combined informational events, Tutorial Activities and System Descriptions, were by far the most frequently occurring events during the interview and each occurred approximately the same number of times.

Transition matrices were constructed for all of these major groupings of events so that a transition pattern could be established. That is, the frequency with which certain events were either preceded by or followed by other events. However, the most obvious observation about these matrices was their nondeterministic nature, that is the dispersal of transitions

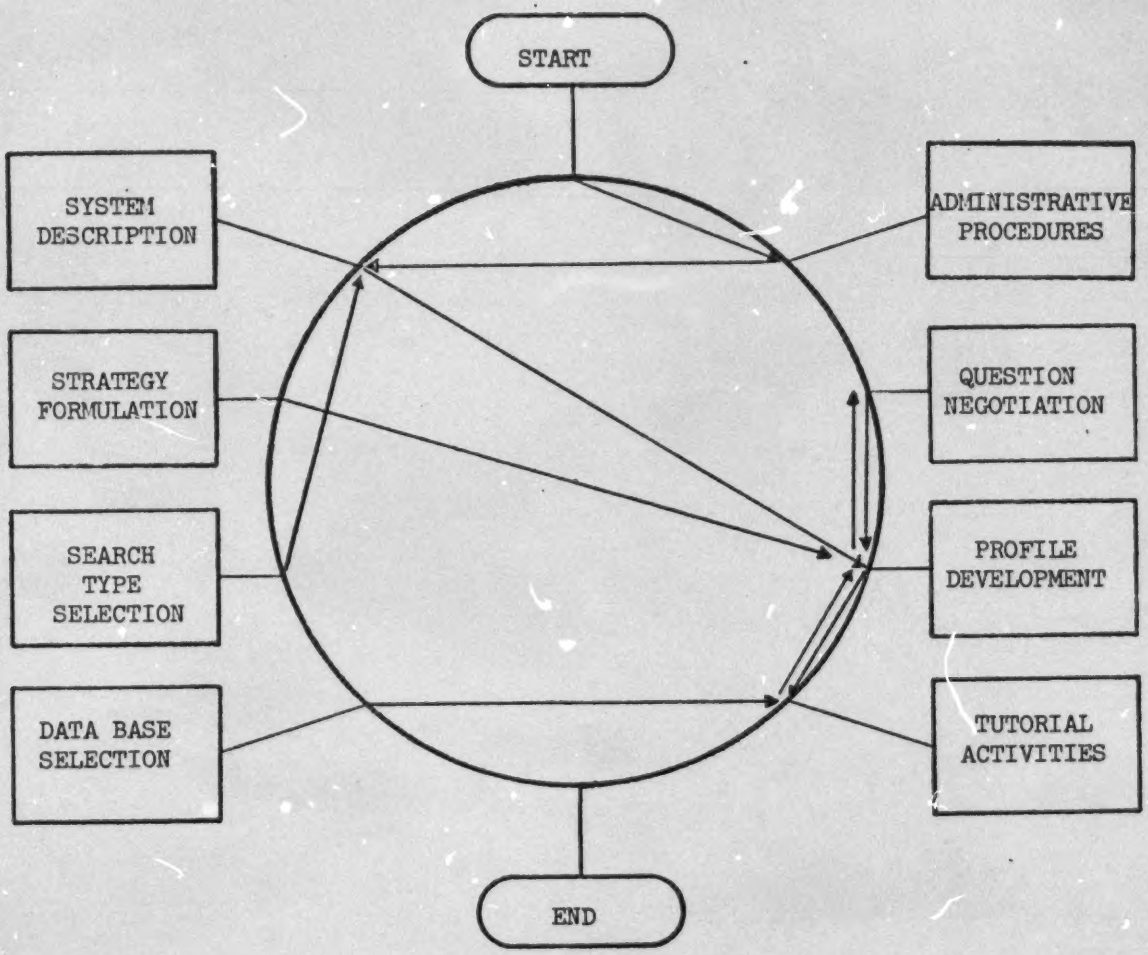
rather than the concentration of transitions. The dispersals were such that graphic representations are difficult to prepare since virtually all connections between the major groupings occurred. The next slide, which corresponds to the matrices is a graphic presentation of the event paths and shows the most frequent exit from each grouping. The double arrows between Profile Development and the two adjacent groupings indicate the strong interactions which are occurring between these three groupings of events. We will see that Strategy Formulation is also very interactive with this group of events.

This difficulty in predicting event transitions was also experienced by the UCLA group when they conducted their event analysis tasks. During the initial stages of the analysis, the researchers used transition matrices to locate occurrences of patterns of event sequences, and concluded that the relationship between events were quite weak. Just as at Georgia, the UCLA group found that Profile Development, Tutorial Activities, and Question Negotiation, were the most frequently occurring events. The UCLA data showed that the relationships between the other major categories and Profile Development indicated that the interview process centered on Profile Development and that the overall process is not a linear one.

So at both sites, the event analysis showed that the interview process is not a linear process nor is it one which has distinct event patterns. The next slide shows the interview event flow diagram developed by UCLA and dubbed the "Spaghetti Model" for obvious reasons. The UCLA model demonstrates the complexity of the process. Each of the circles represent unique events and

abbreviations are given just to show the iterative nature of question negotiation, profile development, tutorial activities and strategy formulation which is illustrated by the number of arrows showing entry to these events from other events. The "spaghetti effect" was attributed to the fact that the interview interaction is a flexible process which takes its structure from the unique characteristics of each intermediary, user, and search request.

Another way of looking at the event analysis data is by the interview time period during which the events occurred. The time at which each event began was tabulated by quartile period and the next slide shows this quartile data for the two major groupings Question Negotiation and Search Strategy Formulation. They are plotted here as percentage of occurrences. For example, 57% of Question Negotiation activities occurred in the first quarter of the interviews and then falls off through the remaining three quartiles. The beginning of Question Negotiation early in the interview is not unexpected, but the fact that 40% of the question negotiation continues into the last three time divisions suggests that this process is in many cases repetitive and also the Search Strategy Formulation events closely parallel the Question Negotiation events suggesting that these two groupings may be closely interrelated, rather than occurring widely separated as is suggested in most linear profile coding models. I should point out that these two cross over in the second quartile period.


START
SYSTEM DESCRIPTION
STRATEGY FORMULATION
SEARCH TYPE SELECTION
DATA BASE SELECTION
ADMINISTRATIVE PROCEDURES
QUESTION NEGOTIATION
PROFILE DEVELOPMENT
TUTORIAL ACTIVITIES
END

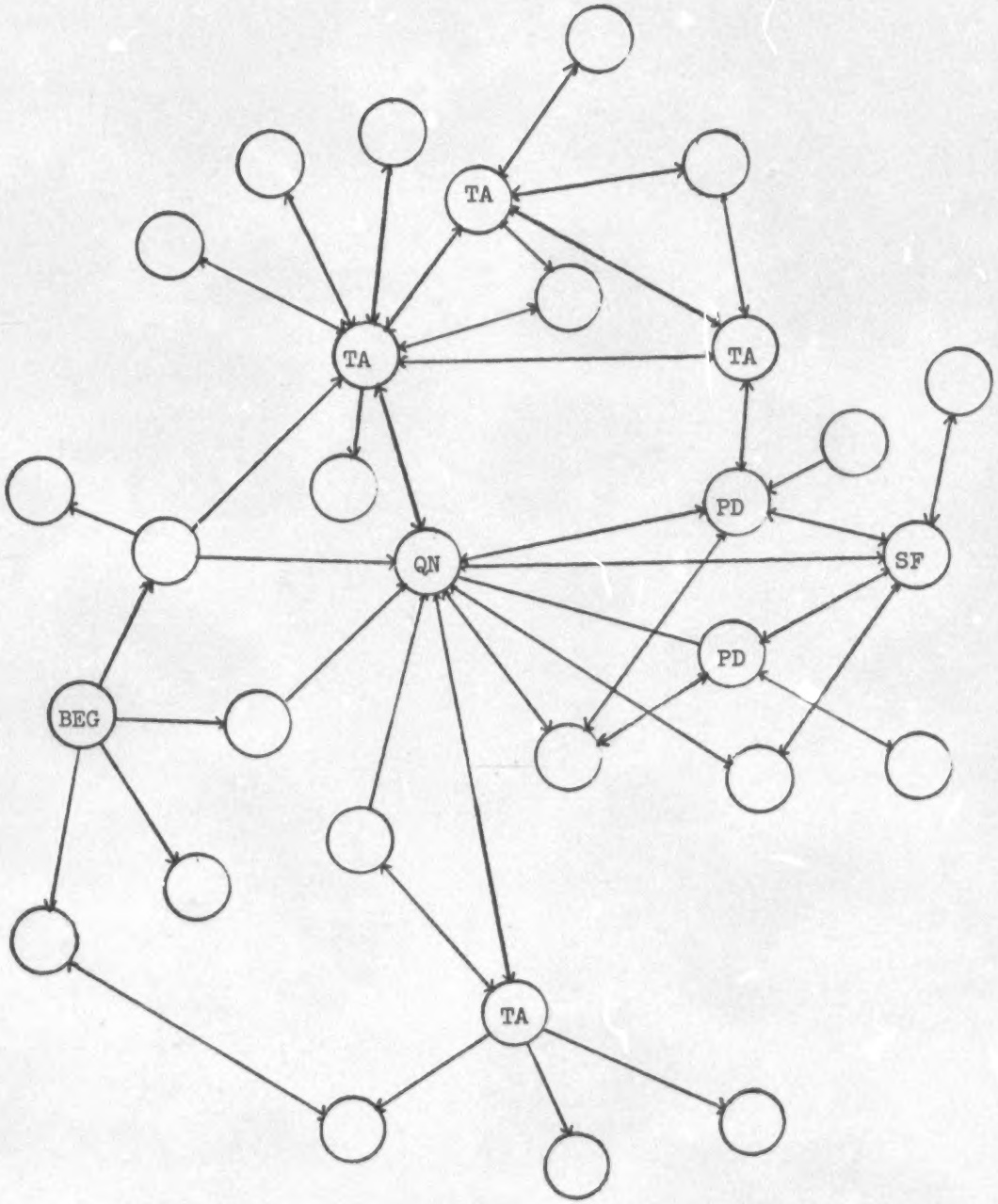
Slide 2

The next slide shows that the Profile Development and Tutorial groupings are also roughly parallel having the highest frequencies in the second and third quartile time periods. The time series data on this slide and the previous slide supports the transition matrices data in showing strong interactions between these group pairs.

So, based on the results from both centers linear models of computer based retrieval operations are viewed as poor representations of the actual process. And, to the extent that they have served as models for training intermediaries, they may also be misleading. This study concluded that there are three predominant characteristics of the user-intermediary interaction process which have major implications for a more representative model. As shown on the next slide these characteristics are, that the process is nondeterministic, it is highly adaptive, and it has a large instructional component.

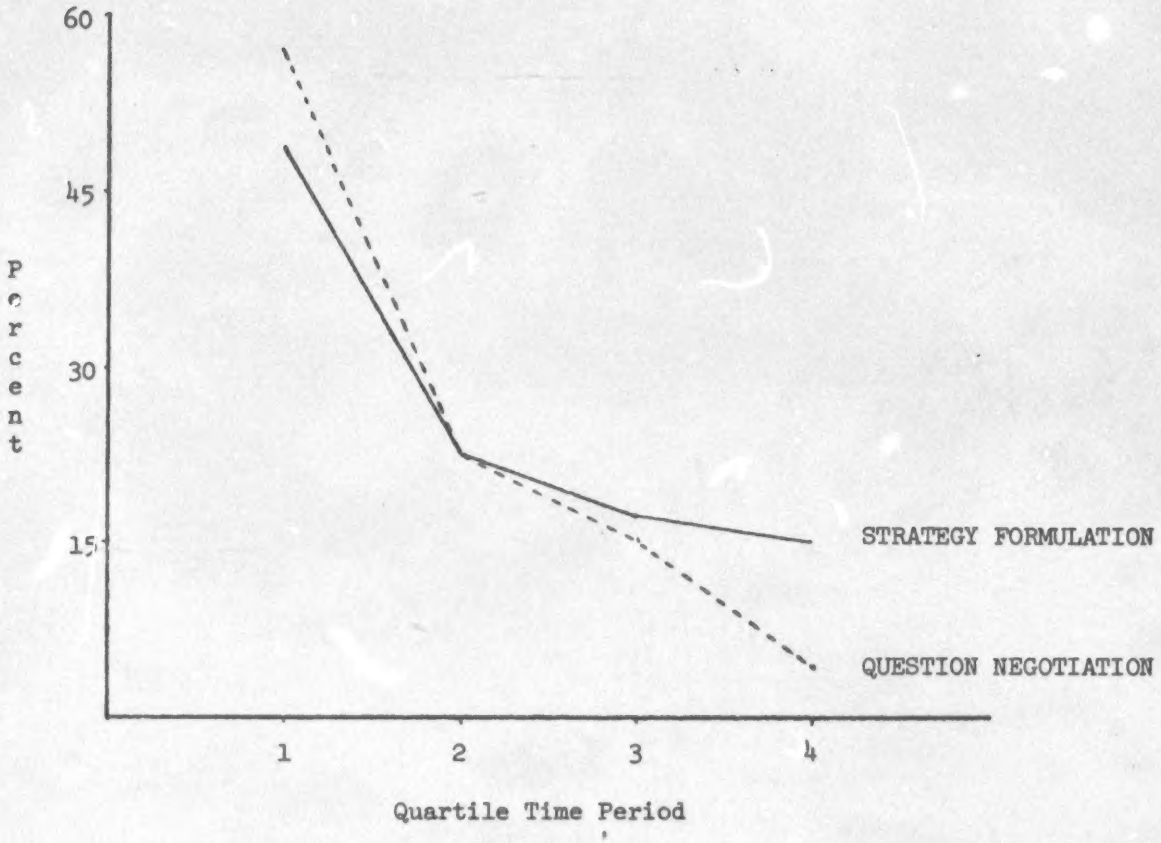
The nondeterministic nature of the process has been shown by the time and sequence analyses of the events taking place in the pre-search interview with the user and, to a lesser extent, in post-search events associated with review and distribution of the search results. Very few of the user and question characteristics have predictive value for the characteristics of the profile, the interaction with the intermediary, or the search results.

Secondly, the use of human intermediaries makes the existing interactive process highly adaptive to perceived needs and preferences of the individual users. Intermediaries obviously respond to a much wider range of stimuli from the user than simply

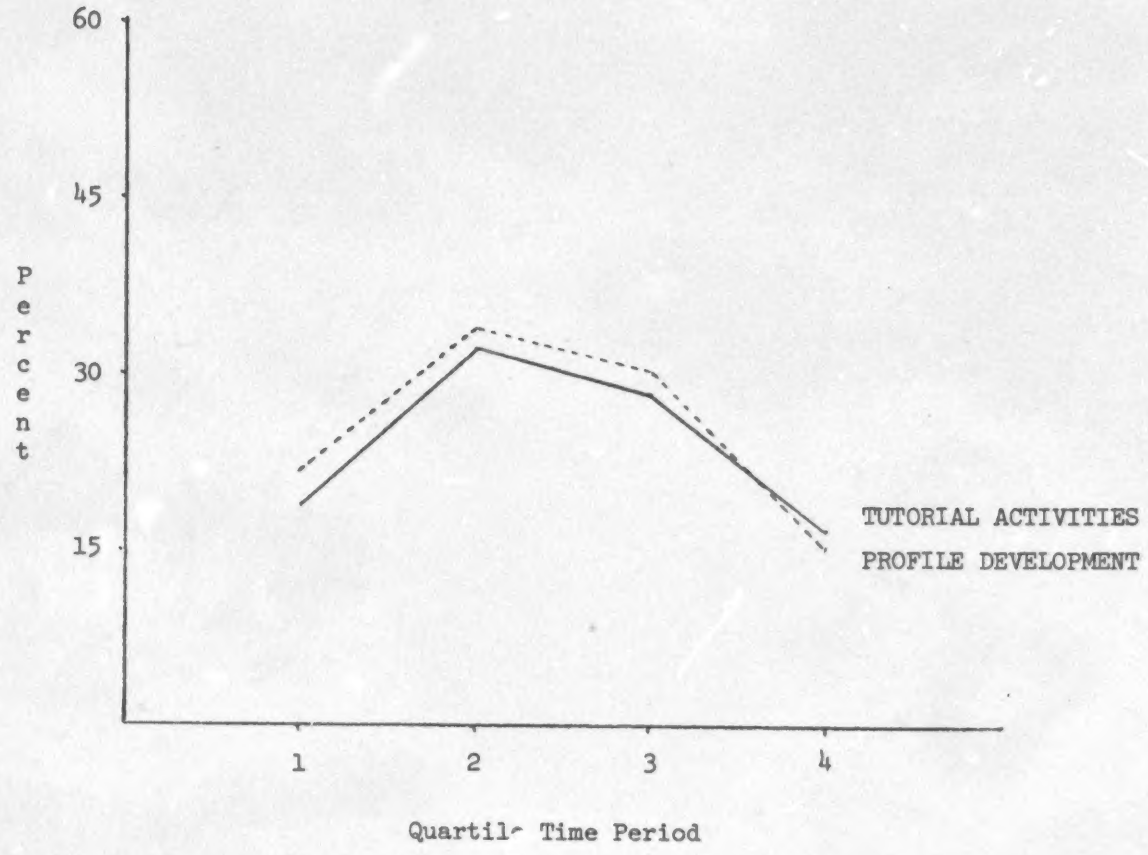


"SPAGHETTI MODEL"

Slide 3


60
45
30
15
Percent
STRATEGY FORMULATION
QUESTION NEGOTIATION
1
2
3
4
Quartile Time Period

Slide 4


60
45
30
15
Percent
1
2
3
4
Quartil- Time Period
TUTORIAL ACTIVITIES
PROFILE DEVELOPMENT

Slide 5

NONDETERMINISTIC PROCESS

HIGHLY ADAPTIVE PROCESS

LARGE INSTRUCTIONAL COMPONENT

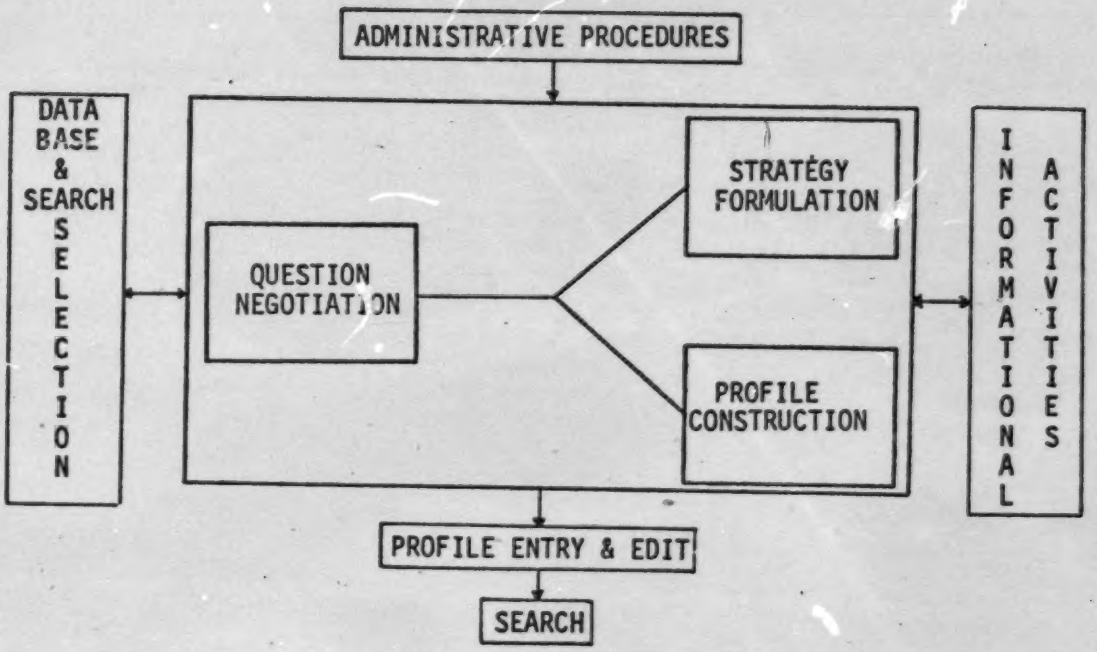
Slide 6

explicit verbal communication. Supportive social and psychlogical activities are provided for the inexperienced or unsure users while interactions with experienced users are optimized toward an efficient and cooperative joint effort.

And thirdly, there appears to be a substantial instructional role being assumed by the intermediaries even though it is not necessarily perceived as "education" by either the intermediary or the user. Its effects are demonstrated through improved search performance and greater user satisfaction as experience is gained with the computer based services. The most critical instructional or advisory roles appear to be related to the vocabulary, subject coverage, indexing conventions, and data content of the data bases, rather than conventions of the retrieval system itself.

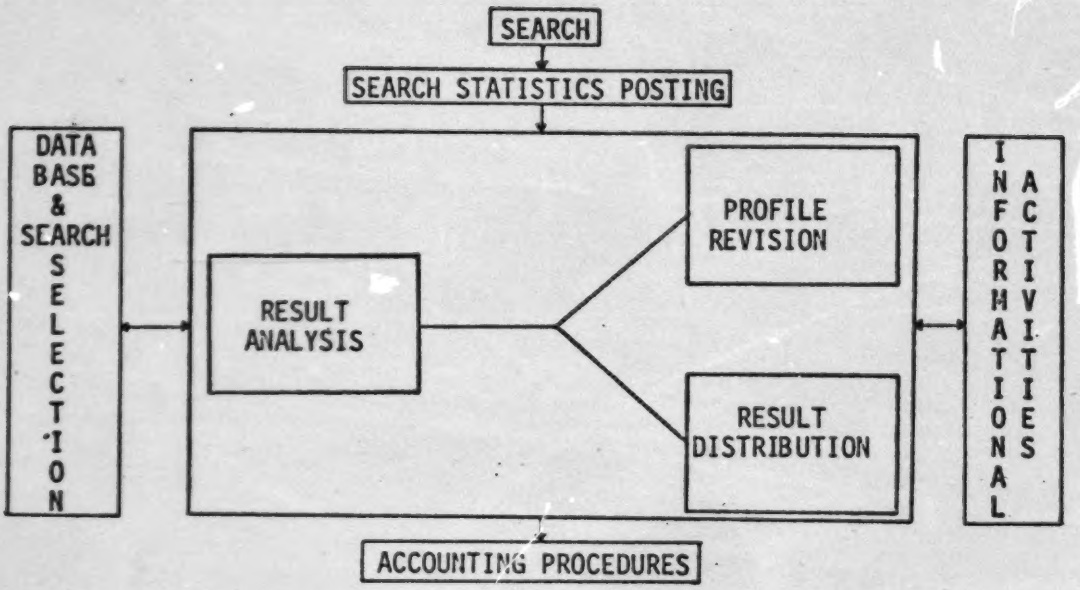
The next slide is a schematic model which is proposed as representative of the existing processes up to the point of search. Within this frame work, the model indicates a symmetric pattern of human mediated activities prior to the search. The three major functions of question negotiation, profile development, and strategy formulation are closely related to each other, but all possible paths of interaction with other events occur.

The next slide represents the proposed model for post-search processes. Here again a symmetric pattern of activities are indicated which involve human mediation. As in the previous slide, no connections are shown for the three central events since they may occur in any combination and some may not occur at all.



PRE-SEARCH MODEL OF EXISTING INTERFACE

Slide 7



POST-SEARCH MODEL OF EXISTING INTERFACE

Slide 8

What, then, are the implications of the results of this study of the existing interface and the proposed model for a network environment? In a network configuration which has one standard interface and encompasses all search centers, the study results indicate that such a network interface must have the five characteristics shown on the next slide:

1. Based on the results of this study, we feel that a network interface should include extensive support for the vocabulary, indexing policies, and document coverage characteristics of the various data bases. This support is needed not only for the new or occasional user but also for the professional intermediaries who cannot possibly know in detail the full range of available data bases. Even in centers where intermediaries tend to specialize in certain of the data bases, it is still necessary for them to use supplemental reference materials in a large majority of the cases. In a network environment which allows direct interaction of the users with the system, it is unrealistic to think each user will have the full range of necessary resources readily available and will take the time and effort to become proficient in their use. Consequently, such support must be provided as part of the retrieval system itself.
2. Also, a network model must accommodate a wide range of users, including the completely inexperienced undergraduate or graduate student, and experienced users who have used the retrieval services a number of times, as well as expert intermediaries who use the system daily in support of their

DATA BASE SUPPORT

WIDE RANGE OF USERS

HUMAN INTERMEDIARIES

DIAGNOSTIC CAPABILITY

MULTIPLE INTERACTION MODES

Slide 9

activities. Translated into system design specification, this means that the system should have something like a "tutorial language" mode for the new and inexperienced user, a "prompt" mode for the occasional user who doesn't use it often enough to remember the commands and messages but would be irritated by a slow tutorial approach, and a straight forward "data entry" mode for the expert user who neither needs nor wants assistance. That is, any given system should have multiple user access modes.

3. Thirdly, the results of this study are believed to support the position that human intermediaries must be retained as an integral component of a network model, though not necessarily functioning in the same way as they do now. Specifically, the human element is critical for the problem solving activities. How much human mediation will be required will depend entirely on the effectiveness of any automated diagnostic capabilities. Programmed user assistance is bound to become more effective as time goes by, but at least at this point it is not conceivable that wide and successful usage of a search system can exist without the human intermediary.
4. Fourth, if a model is assumed in which a user has direct interaction with the retrieval system, then extensive diagnostic or error detecting mechanisms are needed to protect both the user and the system. Reasonable checks must be built into the system to detect when the user is in trouble, even though he may not recognize it himself. It is doubtful, however, that any automated system can ever approximate the

human intermediaries' facility for knowing the user means "no" or "maybe" when he says "yes".

5. And finally, we feel that alternative interaction mechanisms for the user should be available. Although many users would be willing or even prefer to prepare their own profiles, there is still a sizeable group of users who do not want to do so. It also seems reasonable that users' attitudes will change with time and experience, implying a change in their preferences for interaction mechanisms. It is also possible that the choice is a function of the search request, a given user at one time preferring direct interaction for a well defined question, for example, and at another time seeking the assistance of an intermediary to help refine the actual information needs. While a "take it or leave it" approach can be argued, it seems preferable to include the full range of desirable interaction mechanisms in the network model itself, and impose constraints only where practical implementation considerations make them necessary.

In summary, I would like to restate the major findings of this study and that is that the existing model produced as a result of this research does not agree with most of the literature published in this area, which tends to represent the user interface as a linear sequence of activities. This study has found that the interface for the Georgia and UCLA centers examined is not linear, it is highly adaptive, and it has an important instructional component. Another significant finding of the study is that the human intermediary is an integral part of the interface and is essential to the adaptive capability of the interface.